To: Lower Willamette Group

From: EPA

Date: October 26, 2005

Subject: Benthic Interpretive Approach for Portland Harbor Ecological Risk Assessment

This memorandum is in response to LWG's July 11, 2005 presentation on a proposed benthic toxicity interpretation approach for the Lower Willamette River/Portland Harbor and subsequent sharing of information. EPA and its partners appreciate LWG's efforts in preparing the presentation. EPA also appreciates the LWG's recent quick response to our request for a working copy of your power analysis for the benthic predictive model. This serves as a significant example of progress towards a collaborative approach for assessing risk at the Portland Harbor site.

This memo describes EPA's expectations regarding the predictive benthic toxicity modeling and analysis for the Portland Harbor site. It is critical that we reach agreement on key elements of the benthic toxicity approach such as the toxicity hit/no-hit determination. This memo also provides direction on other aspects of the benthic interpretive approach for Portland Harbor, including the floating percentile methodology, reporting, detections for inclusion in the model, data qualifiers, data chemistry preparations, and deliverables.

EPA would like to meet with the LWG to discuss and reach agreement on the items outlined below within the next 30 days. This will necessitate postponing submittal of the benthic toxicity interpretation report until mid-December or early January.

Hit/No Hit Identification

Pooled Endpoints: Toxicity test evaluations should consider the combined results of growth and survival for each test species (i.e., a sample should be considered toxic if the results for either survival or growth are determined to be toxic), as well as the results of both tests combined rather than evaluating each test endpoint separately – *Hyalella* growth and mortality and *Chironomid* growth and mortality. While evaluating the two survival endpoints separately is acceptable, we would place little weight in looking at growth as separate endpoints, due to the confounding effects between mortality and growth. EPA understands that the LWG is combining growth and survival into a single result for each species and, in addition, combining the *Hyalella* and *Chironomus* results. EPA is supportive of this approach.

Control-Normalization: Control-normalized responses for survival and growth endpoints should be calculated as test response/control response (T/C). Using a control-normalized approach provides more consistency in comparisons among batches and does not give additional slack to tests with poor control performance.

Hit/No Hit Toxicity Thresholds:

Please include the following toxicity thresholds in your evaluation of predictive models:

1) < 90% control-normalized survival OR < 90% control-normalized growth

2) < 80% control-normalized survival OR <80% control-normalized growth

3) < 70% control-normalized survival OR <70% control-normalized growth

Only samples statistically different from controls should be considered to be below the toxicity threshold. Samples below the threshold and not statistically different from control should be evaluated for adequate power. Samples with insufficient power should be excluded from the analyses.

Toxicity thresholds should be evaluated for the combined *Hyalella* survival and growth, the combined *Chironomus* survival and growth, and the results of both tests combined.

EPA and its partners believe that, because most results fall between 10 and 30 percent, adding an additional level there rather than at 50 percent, as proposed by LWG, will provide for a better analysis of magnitude of effects.

Floating Percentile Methodology

The government team is currently reviewing the use of the floating percentile method (FPM) in analyzing Portland Harbor benthic toxicity and chemistry data. EPA believes open communication with LWG will help to resolve questions that we have about the FPM and user specificity prior to LWG's formal submittal of results based on this approach. Because the FPM does not necessarily yield a unique solution, different users may derive a range of results from the same data. In order for EPA and its partners to determine whether FPM results are reproducible and acceptable, it is critical that our users have sufficiently detailed instructions on the use of the FPM. To date, technical staff at the Oregon Department of Environmental Quality (DEQ) has had some difficulty in reproducing results, but we believe these issues can be resolved through collaboration. Progress would be expedited if technical staff at DEQ could communicate directly and as soon as possible with LWG's consultants regarding questions about the FPM.

More specifically, DEQ has been able to reproduce the initial spreadsheet results of the LWG's FPM approach using the provided instructions. However, some of the final results vary by orders of magnitude. DEQ believes this can be addressed by programming the manual part of

the calculations into the spreadsheet to provide consistent application of the process. Additional issues raised by DEQ's attempts to replicate the LWG's FPM results include the following:

- The instructions provided on how to establish a percent tolerance for false negatives within the FPM apparently eliminate the option of an agency selecting their own percent tolerance for false negatives.
- In its trial runs, DEQ is producing unexplained results at various selected percentages of false negatives.
- Some of the LWG results using the FPM appear to be inconsistent with the established conditions for running the model.

More detail is available on these issues. We are hopeful that they can be resolved through direct, open and expeditious communication at the appropriate levels.

Reporting

Please provide an electronic data table listing the statistical results for all samples for each toxicity test endpoint, including an identification of samples with insufficient statistical power. Please also provide a detailed description of the approach used to determine statistical significance.

Detections for Inclusion in the Model

The LWG has proposed a minimum of 30 detections before a contaminant is included in the model. This number may be acceptable, but additional information should be provided on the contaminants dropped from the analysis to better understand patterns in toxicity separate or in conjunction with modeling efforts. Additional information on these chemicals should be provided including:

- A list of contaminants detected in sediment, but not included in the model (detected fewer than 30 times).
- Map areas where these chemicals were detected to provide information on where these detections are in the harbor.
- Discussion of physical properties (e.g. VOCs) – this may link with groundwater plumes or other lines of evidence available.
- Potential correlation with false positive results from modeling effort.

Data Qualifiers

EPA seeks clarification on the "N" qualifier. Specifically, what does this mean and is it appropriate to remove these?

Data Chemistry Preparation

TPH: It is unclear how TPH is being handled in the model. We need to understand whether TPA was analyzed, and if so, what it is being called in the database. For example, is oil and grease included in the residual range?

Normalization: Normalization of organic chemical concentrations to organic carbon is not required. Consider normalizing to percent fines or developing an alternative approach to evaluating the impact of the amount of fine-grain sediment on the predictive models. A separate model may need to be developed for fine grained sediments, which can then be compared to the larger model and differences can be compared.

Interim Submittals

EPA requests the continued submittal of information as necessary to reach timely resolution of the above issues. Is this a request (i.e., items in this memo), or just a general comment?